

[illegible]

(HSE-) HXSD INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D:  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J:  
 PI Zhao QH, Zhou P, Goodrich R, Dimanac RT:  
 XX WPI: 2001-442253/47.  
 DR P-PSDB: AAM40129.

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX Claim 1: SEQ ID NO 1488; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nucleotide,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilization of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 1557 BP; 581 A; 248 C; 320 G; 408 T; 0 other;

Query Match 43.4%; Score 1460.2; DB 22; Length 1557;

Best Local Similarity 99.4%; Pred. No. 0; Mismatches 8; Index 1; Gaps 1;

Matches 1476; Conservative 0; Mismatches 8; Index 1; Gaps 1;

XX 1882 gattcttggaactcaacctatgaacatgagcattcaaatggaataaagctgcaaaa 1941  
 DB 61 gattttggaactcaacctatgaacatgagcattcaaatggaataaagctgcaaaa 120  
 XX 1942 aagaagaatcgcgaagaacggtttgtgcagaacattggaagaatgacatgagcccta 2001  
 DB 121 gaaggaatacgcggaagaacggtttgtgcagaacattggaagaatgacatgagcccta 180  
 XX 2002 caaatlaaagacacatcgaatgataatgatactgatactcattggaacttccataat 2061  
 DB 181 caaatlaaagacacatcgaatgataatgatactgatactcattggaacttccataat 240  
 XX 2062 gaaggaagaatgaagaagttttgacatcagaatgataatgataatgataatgataat 2121  
 DB 241 gaaggaagaatgaagaagttttgacatcagaatgataatgataatgataatgataat 300  
 XX 2122 gagatctgcatgagtgatgataatgataatgataatgataatgataatgataatgataat 2181  
 DB 301 gagatctgcatgagtgatgataatgataatgataatgataatgataatgataatgataat 360  
 XX 2182 acagata-gattttcagacttattttttgaaacaataaattgttgaagaagctggc 2240  
 DB 361 acagatacagacttattttttgaaacaataaattgttgaagaagctggc 420  
 XX 2241 tgaataaccagataatgaataatgaataatgataatgataatgataatgataatgataat 2300  
 DB 421 tgaataaccagataatgaataatgataatgataatgataatgataatgataatgataat 480  
 XX 2301 atatactagactgaggaatcagcagaggaataatccttcaaaaacagcagagtg 2360  
 DB 481 atatactagactgaggaatcagcagaggaataatccttcaaaaacagcagagtg 540  
 XX 2361 atatgcgtcttccagtgatgataatgataatgataatgataatgataatgataatgataat 2420  
 DB 541 atatgcgtcttccagtgatgataatgataatgataatgataatgataatgataatgataat 600

XX 2421 ccaccatctgattggaactgagacacagcagtgatgataatgataatgataatgataatgataat 2480  
 DB 601 ccaccatctgattggaactgagacacagcagtgatgataatgataatgataatgataatgataat 660  
 XX 2481 aagaagaatcagtaatgataatgataatgataatgataatgataatgataatgataatgataat 2540  
 DB 661 aagaagaatcagtaatgataatgataatgataatgataatgataatgataatgataatgataat 720  
 XX 2541 ggcgaagaagaatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2600  
 DB 721 ggcgaagaagaatgataatgataatgataatgataatgataatgataatgataatgataatgataat 780  
 XX 2601 caatgaaatagcattgctcagcagcctgtgtcgaacagagctgataatgataatgataatgataat 2660  
 DB 781 caatgaaatagcattgctcagcagcctgtgtcgaacagagctgataatgataatgataatgataat 840  
 XX 2661 cctgtgtgtcagcagtgatgataatgataatgataatgataatgataatgataatgataatgataat 2720  
 DB 841 cctgtgtgtcagcagtgatgataatgataatgataatgataatgataatgataatgataatgataat 900  
 XX 2721 gaagatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 2780  
 DB 901 gaagatgataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 960  
 XX 2781 taagatttgggaattacagatgataatgataatgataatgataatgataatgataatgataatgataat 2840  
 DB 961 taagatttgggaattacagatgataatgataatgataatgataatgataatgataatgataatgataat 1020  
 XX 2841 tatgtcgaagaatcagataatgataatgataatgataatgataatgataatgataatgataatgataat 2900  
 DB 1021 tatgtcgaagaatcagataatgataatgataatgataatgataatgataatgataatgataatgataat 1080  
 XX 2901 tgtgtacactgtgtcgtggaagaatgataatgataatgataatgataatgataatgataatgataatgataat 2960  
 DB 1081 tgtgtacactgtgtcgtggaagaatgataatgataatgataatgataatgataatgataatgataatgataat 1140  
 XX 2961 gacccagaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 3020  
 DB 1141 gacccagaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 1200  
 XX 3021 tgcgcactatcaataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 3080  
 DB 1201 tgcgcactatcaataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1260  
 XX 3081 gatgtgcagaagaatgataatgataatgataatgataatgataatgataatgataatgataatgataat 3140  
 DB 1261 gatgtgcagaagaatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1320  
 XX 3141 caaataatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 3200  
 DB 1321 caaataatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 1380  
 XX 3201 caatctgacatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 3260  
 DB 1381 caatctgacatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 1440  
 XX 3261 tttctttaaataatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 3320  
 DB 1441 tttctttaaataatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaatcaggaat 1500  
 XX 3321 ctacagaactgacataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 3365  
 DB 1501 ctacagaactgacataatgataatgataatgataatgataatgataatgataatgataatgataatgataat 1545

## RESULT 2

AA161071

ID AA161071 standard; cDNA; 1382 BP.

XX AC AA161071:

XX 22-OCT-2001 (first entry)

[illegible]

QY	2122	gaqatitgibatigigtatgaaagttgagatgatatlaaagaaccccttgaacttgatgaa	2181
Db	133	gaqatitgibatigigtatgaaacatgagagatattlaaagaaccccttgaacttgatgaa	192
QY	2182	acaagata-qatitctcatgacttattllttgaaacaataaat-gttagaagagcttgg	2239
Db	193	acagatgacgatctctcatgacttattllttgaaacaataaatagtgtgaatgagcttgg	252
QY	2240	ctgaaaccccgagaatatgaaatgaaagctgacccaattaaqaaataccataatggac	2299
Db	253	ctgaaaccccgataatgaaatagaaagctgacccaattaaqaaataccataatggac	312
QY	2300	aatatctagagactgaggaatctgcacgcagaggaataatctttaacaacaacgacagagtg	2359
Db	313	aatactagagactgaggaatctgcacgcagaggaataatctttaacaacaacgacagagtg	372
QY	2360	catatgagcttcccgatgtgatactgaaataagaaatattgtaagtaagatgcaag	2419
Db	373	catatgagcttcccgatgtgatactgaaataagaaatattgtaagtaagatgcaag	432
QY	2420	cccacacatctgattgagctgtgacacacagatgagtttcaaaccccatgacacagaatgac	2479
Db	433	cccacacatctgattgagctgtgacacacagatgagtttcaaaccccatgacacagaatgac	492
QY	2480	aaaaagaatctatagttaaatcttcgcacatcgagaataatctatctgtcttactgctarcacag	2539
Db	493	aaaaagaatctatagttaaatcttcgcacatcgagaataataatctgtcttactgctarcacag	552
QY	2540	tggcgaagaagagcttctgatattaaagaatgtaaacattgtatccgttatgtctgtca	2599
Db	553	tggcgaagaagagcttctgatattaaagaatgtaaacattgtatccgttatgtctgtca	612
QY	2600	ccaatgaaatgagccatgagcttcacagcccgctgctcagccagagctgatgagacacactacg	2659
Db	613	ccaatgaaatgagccatgagcttcacagcccgctgctcagccagagctgatgagacacactacg	672
QY	2660	tccctggttgcctcacagtggtctcagagatattcgaacatgagacagttacgattcccgag	2719
Db	673	tccctggttgcctcacagtggtctcagagatattcgaacatgagacagttacgattcccgag	732
QY	2720	agaagaatcatatlaaagctatcacatctggttcacaatatgaaacccgaagagatagctc	2779
Db	733	agaagaatcatatlaaagctatcacatctggttcacaatatgaaacccgaagagatagctc	792
QY	2780	alaagatctttagaatctacagatgcgaagatgataatgtgaaagaagaatgaaacccaagagaa	2839
Db	793	alaagatctttagaatctacagatgcgaagatgataatgtgaaagaagaatgaaacccaagagaa	852
QY	2840	atatgtccgaagcattacagaataaaccccatcactaatacttcccttgcgaaaaactgcga	2899
Db	853	atatgtccgaagcattacagaataaaccccatcactaatacttcccttgcgaaaaactgcga	912
QY	2900	gtggtgcttagcctgtctcgtggggaagatatcctatgtattagaataatgcatcacgtcataa	2959
Db	913	gtggtgcttagcctgtctcgtggggaagatatcctatgtattagaataatgcatcacgtcataa	972
QY	2960	tgaccccggaattccaagaactttacatctgtaagagtaaaacaacacactgcgaaaaagaagt	3019
Db	973	tgaccccggaattccaagaactttacatctgtaagagtaaaacaacacactgcgaaaaagaagt	1032
QY	3020	gtgcgcgactatcaataaatggttgaatactctgcacaatgtgcccagqcttggaggaaca	3079
Db	1033	gtgcgcgactatcaataaatggttgaatactctgcacaatgtgcccagqcttggaggaaca	1092
QY	3080	tgaatgtgacaaaagctttagatttgccttgcctcctaataatagqaattttttagttagtct	3139
Db	1093	tgaatgtgacaaaagctttagatttgccttgcctcctaataatagqaattttttagttagtct	1152
QY	3140	tcaaaaataattccaacaagaagaacaatacacaanaagtggtgtagaatctacccatcacattc	3199
Db	1153	tcaaaaataattccaacaagaagaacaatacacaanaagtggtgtagaatctacccatcacattc	1212
QY	3200	ccaactcttgactatccgaatgctgttattatgtagtaggaattagcacttgatttgaag	3259

Dd	1213	cgaatcttgcatcaccgaatgctgtttatttcgatgatgagatttaacattgtagaa	1272
Gy	3260	attctttaaaactatcgcttcaaacattttaataatgatatgatatattt-catta	3318
Dd	1273	attctttaaaaataactatcagcttcaaacattttaataatgatatgatatatttccataa	1332
Gy	3319	tgcataagaactgacataagaatcataaataatattgttttacctg	3365
Dd	1333	tgcataagaactgacataagaatcataaataatattgttttacctg	1379
RESULT	3		
ID	AAS01149	standard; DNA; 1258 BP.	
XX	AAS01149;		
XX	12-SEP-2001	(first entry)	
XX		Interferon induced nucleic acid, IFN4.	
KM		Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;	
KM		immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;	
KM		graft rejection; viral infection; hepatitis; aplastic anemia; cancer;	
KM		human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;	
KM		hematologic disease; chronic neutropenia; myocardial infarction;	
KM		neurological disease; Alzheimer's disease; Parkinson's disease; tumour;	
XX		amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN4; ds.	
XX	Homo sapiens.		
OS			
FH	Key	Location/Qualifiers	
FH	CDS	1..915	
FT		/*tag-a	
FT		/product-"IFN4"	
FT		/note-"Interferon induced polypeptide"	
XX	MO200118208-A2.		
FN			
Pd	15-MAR-2001.		
PE	08-SEP-2000; 200OMO-US24704.		
PR	08-SEP-1999; 99US-0152921.		
PR	20-OCT-1999; 99US-0160575.		
PR	20-JAN-2000; 200OUS-017104.		
PR	07-SEP-2000; 200OUS-0656633.		
CC	(CURA.) CTRAGEN CORP.		
CC	(BIOJ ) BIOGEN INC.		
XX			
Pj	Peyman JA, Da Silva A, Hochman P, Hsu AJ;		
DR	WPI: 2001-235201/24.		
DR	P-PSDB: AAU00296.		
PT	New interferon induced polypeptides and polynucleotides, useful for the		
PT	diagnosis, prevention and treatment of immunological, cell		
PT	proliferative disorders, such as lupus erythematosus, cancer, stroke		
PT	and Alzheimer's disease -		
PS	Claim 9; Page 29-32; 134pp; English.		
CC	The sequence represents interferon induced nucleic acid, IFN4. IFN		
CC	nucleic acids and polypeptides are useful for treating or preventing a		
CC	pathology associated with IFN polypeptide in a human. They are useful for		
CC	determining the presence of or predisposition to a disease associated		
CC	with altered levels of IFN polypeptide or polynucleotide. IFN nucleic		
CC	acids, polypeptides and antibodies are useful for diagnosis, prevention		
CC	or treatment of variety of immunological and cell proliferative		
CC	disorders, such as autoimmune diseases e.g. lupus erythematosus,		
CC	immunodeficiency diseases such as acquired immunodeficiency syndrome		

CC (AIDS), graft rejection, viral infections including hepatitis and human  
CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,  
CC hematologic diseases such as aplastic anaemia and chronic neutropenia  
CC and cancer. In addition they are also useful for treating or diagnosing  
CC various disorders associated with cell death, including myocardial  
CC infarction, stroke, neurological diseases including Alzheimer's and  
CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular  
CC atrophy. IFN nucleic acids and polypeptides are also useful for  
CC identifying interferon-like proteins and interferon agonists, for  
CC screening drugs and compounds which inhibit or enhance IFN activity or  
CC function and as targets for the identification of small molecules that  
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal  
CC cell or tumour cell growth in mammals, including humans.

Query Match	23.9%	Score 803;	DB 22;	length 1256;
Best Local Similarity	95.5%;	Pred. No. 3.8e-189;		
Matches 853; Conservative	0;	Mismatches 0;	Indels 40;	Caps 1

QY	45	gcaacagagcgttgaacccctctgtctctctaaagtggcgacgagcaacgagcagacgaatttca	104
Db	55	gcacacagagcgttgaacccctgtctctctctaaagtggcgacgagcaacgagcaacgaatttca	114
QY	105	ccctgtcccgagacaacagcaacacatctgtcttgagagaacccctctccctctctgaagaag	164
Db	115	ccctgtcccgagacaacagcaacacacatctgtcttgagagaacccctctccctctctgaagaag	174
QY	165	aaagatgtcgaattgagatattccacaacaggaattctccgtatctatctccgtttcag	224
Db	175	aaagatgtcgaattgagatattccacaacaggaattctccgtatctatctccgtttcag	234
QY	225	ggcccaagggttgaanaatgtacatccagcttgagacccctgtcttgactaccttgacctctgc	284
Db	235	ggcccaagggttgaanaatgtacatccagcttgagacccctgtcttgactaccttgacctctgc	296
QY	285	tgcagagagtgaaagagcagatctcagagagacagctccgccacctctcggagacaatgcagagat	344
Db	257	--cagagagtgaaagagcagatctcagagagacagctccgccacctctcggagacaatgcagagat	314
QY	345	tgaactgtgtctgagacactttggagaaaggaatcttgacacttggttggatcttgggaatt	404
Db	315	tgaactgtgtctgagacactttggagaaaggaatcttgacacttggttggatcttgggaatt	374
QY	405	cgttgaagccctctcgagaaacccgacagccctctgcgcgcgcgtctacatgaacctgagct	464
Db	375	cgttgaagccctctcgagaaacccgacagccctctgcgcgcgcgtctacatgaacctgagct	434
QY	465	caaggaacttgcctctccatcgttttgaagaagctcatgatgaatlatctccaacttgcgaa	524
Db	435	caaggaacttgcctctccatcgttttgaagaagctcatgatgaatlatctccaacttgcgaa	494
QY	525	ccctctctcagcccactcttgttgagcaagcttctcagttagagacgctcttgataaagtcat	584
Db	495	ccctctctcagcccactcttgttgagcaagcttctcagttagagacgctcttgataaagtcat	554
QY	585	ggaagagaagaaacttgttgaacaattgagacagaaacccgatttgcgtctgcagaaacaatg	644
Db	555	ggaagagaagaaacttgttgaacaattgagacagaaacccgatttgcgtctgcagaaacaatg	614
QY	645	aaatgaatcaggtgttgaagagagctactaaanaagatgtgtgcagaaagaaacrtggtctc	704
Db	615	aaatgaatcaggtgttgaagagagctactaaanaagatgtgtgcagaaagaaacrtggtctc	674
QY	705	tgcattctctgaagtctctctgttcaaacacagaaacaaatgaactgttctcctaaagttaaaag	764
Db	675	tgcattctctgaagtctctctgttcaaacacagaaacaaatgaactgttctcctaaagttaaaag	734
QY	765	ctctgatttgcacaaagacaaatgacaggaatttgagaaattatccaaagtgtatggtctca	824
Db	735	ctctgatttgcacaaagacaaatgacaggaatttgagaaattatccaaagtgtatggtctca	794

QY 825 agtgaagagcaactcttccacacagctcagcacaatctgagagagctcgtgag 884  
|||||  
DB 795 agtgaagagcaactcttccacacagctcagcacaatctgagagagctcgtgag 854  
|||||  
QY 885 catgagagataactcattcagaatctcttgcagatctctcgtgagttcag 937  
|||||  
DB 855 catgagagataactcattcagaatctcttgcagatctctcgtgagttcag 907  
|||||  
RESULT 4  
AAS01151  
ID AAS01151 standard; DNA; 1270 BP.  
XX  
AC AAS01151;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Interferon induced nucleic acid, IFN6.  
XX  
KW Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;  
immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;  
graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;  
human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;  
haematologic disease; chronic neutropenia; myocardial infarction;  
neurological disease; Alzheimer's disease; Parkinson's disease; tumour;  
amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN6; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key location/Qualifiers  
FT CDS 1..927  
FT /tag= a  
FT /product= "IFN6"  
FT /note= "Interferon induced polypeptide"  
XX  
PM WO200118208-A2.  
XX  
PD 15-MAR-2001.  
XX  
PF 08-SEP-2000; 2000WO-US24704.  
XX  
PR 08-SEP-1999; 99US-0152921.  
PR 20-OCT-1999; 99US-0160575.  
PR 20-JAN-2000; 2000US-0177104.  
PR 07-SEP-2000; 2000US-0656633.  
XX  
PA (CURA-) CURAGEN CORP.  
(BIOJ) BIOGEN INC.  
XX  
PE Peyman JA, Da Silva A, Hochman P, Hsu A;  
XX  
DR WPI: 2001-235201/24.  
DR P-PSDB: AAU000298.  
XX  
PT New interferon induced polypeptides and polynucleotides, useful for the  
diagnosis, prevention and treatment of immunological, cell  
proliferative disorders, such as lupus erythematosus, cancer, stroke  
and Alzheimer's disease  
XX  
PS Claim 9; Page 40-43; 134pp; English.  
XX  
CC The sequence represents interferon induced nucleic acid, IFN6. IFN  
nucleic acids and polypeptides are useful for treating or preventing a  
pathology associated with IFN polypeptide in a human. They are useful for  
determining the presence of or predisposition to a disease associated  
with altered levels of IFN polypeptide or polynucleotide. IFN nucleic  
acids, polypeptides and antibodies are useful for diagnosis, prevention  
or treatment of variety of immunological and cell proliferative  
disorders, such as autoimmune diseases e.g. lupus erythematosus,  
immunodeficiency diseases such as acquired immunodeficiency syndrome  
(AIDS), graft rejection, viral infections including hepatitis and human  
immunodeficiency virus (HIV), immune-mediated glomerulonephritis,  
haematologic diseases such as aplastic anaemia and chronic neutropenia

CC and cancer. In addition they are also useful for treating or diagnosing  
CC various disorders associated with cell death, including myocardial  
CC infarction, stroke, neurological diseases including Alzheimer's and  
CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular  
CC atrophy. IFN nucleic acids and polypeptides are also useful for  
CC identifying interferon-like proteins and interferon agonists, for  
CC screening drugs and compounds which inhibit or enhance IFN activity or  
CC function and as targets for the identification of small molecules that  
CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal  
CC cell or tumour cell growth in mammals, including humans.  
XX  
SQ Sequence 1270 BP; 401 A; 271 C; 279 G; 319 T; 0 other;  
Query Match 23.9%; Score 803; DB 22; Length 1270;  
Best Local Similarity 95.5%; Pred. No. 3.8e-189;  
Matches 853; Conservative 0; Mismatches 0; Indels 40; Gaps 1;  
QY 45 gcaagagcgtagacccctgtctcttaagtggcagcagcagcagcagcattca 104  
|||||  
DB 67 gcaagagcgtagacccctgtctcttaagtggcagcagcagcagcagcattca 126  
|||||  
QY 105 cctgtccgcagacaacagacacacatctgtctgggagacccctcctctcttgagaaag 164  
|||||  
DB 127 cctgtccgcagacaacagacacacatctgtctgggagacccctcctctcttgagaaag 186  
|||||  
QY 165 aaagatgtcgaatgggtatccacagacagaaattccgtatctatctcgtgttcag 224  
|||||  
DB 187 aaagatgtcgaatgggtatccacagacagaaattccgtatctatctcgtgttcag 246  
|||||  
QY 225 ggcagagtgaaatgttaccacagcagtgagcgtgtgtgtgactactgactttcgc 284  
|||||  
DB 247 ggcagagtgaaatgttaccacagcagtgagcgtgtgtgtgactactgactttcgc 268  
|||||  
QY 285 tgcagaggttgaagagcagatctcagagcagtcgcacccctccgagacatgacagcagt 344  
|||||  
DB 269 -cagaggttgaagagcagatctcagagcagtcgcacccctccgagacatgacagcagt 326  
|||||  
QY 345 tgaactgtctgtgacacatttgaagaaaggagctgtgacatttgaactcagaaatt 404  
|||||  
DB 327 tgaactgtctgtgacacatttgaagaaaggagctgtgacatttgaactcagaaatt 386  
|||||  
QY 405 cgtgagagcccttcggaagaaacagcagacccctctgagccgcgcacatgaagccctgagct 464  
|||||  
DB 387 cgtgagagcccttcggaagaaacagcagacccctctgagccgcgcacatgaagccctgagct 446  
|||||  
QY 465 caaggaactgtcctctccatcgtttgagaaagcgcacatgatataatctcaactgtgaa 524  
|||||  
DB 447 caaggaactgtcctctccatcgtttgagaaagcgcacatgatataatctcaactgtgaa 506  
|||||  
QY 525 cctcctcagcccaactcgtgtgagacaagctctcagttagagacgctcttgataagttat 584  
|||||  
DB 507 cctcctcagcccaactcgtgtgagacaagctctcagttagagacgctcttgataagttat 566  
|||||  
QY 585 ggaagagaaactgttgaacaattgaagagaaacagcagattgtctgtcagagaaacaatg 644  
|||||  
DB 567 ggaagagaaactgttgaacaattgaagagaaacagcagattgtctgtcagagaaacaatg 626  
|||||  
QY 645 aatgaatcaagtgtaagaaagctactaaaagaagatttgcaagaaagaaactgtgtctc 704  
|||||  
DB 627 aatgaatcaagtgtaagaaagctactaaaagaagatttgcaagaaagaaactgtgtctc 686  
|||||  
QY 705 tgcatttgaatgttcttcgcaaaacagagaaacaatgaactgttccaagattgaacag 764  
|||||  
DB 687 tgcatttgaatgttcttcgcaaaacagagaaacaatgaactgttccaagattgaacag 746  
|||||  
QY 765 ctctgattgtccagaaagcagatgagagattgaaattatcaacaagtgtgtgtctca 824  
|||||  
DB 747 ctctgattgtccagaaagcagatgagagattgaaattatcaacaagtgtgtgtctca 806  
|||||  
QY 825 agtgaagagcaactcttccacacagctcagcacaatctgagagagcgtcgtgag 884  
|||||  
DB 807 agtgaagagcaactcttccacacagctcagcacaatctgagagagcgtcgtgag 866  
|||||









```

|||||
Db 535 TCGGATATCCAGACAGAGAAATTCGGCTATCTCATCTCGTCTCAGAGGCGAGGTGAA 476
Oy 237 aatgtacatccagtgagtgagctgtgtgtgactgaactgaccttctgctgcagagtgaa 296
Db 475 AATGTACATCCAGAGGTGAGGCGCTGTGCTGACTACCTGACCTTTCCTCGCAGAGGTGAA 416
Oy 297 ggaagacatccagtgagtgagtgagctgtgtgtgactgaactgaccttctgctgcagagtgaa 356
Db 415 GGACGAGATTCAGAGGACAGAGTGCCACCTCCGGGAAACATGAGGAGGATGAACTCTGCT 356
Oy 357 gaggcccttgaggaaggaagtgagtgagctgtgtgtgactgaactgaccttctgctgcagagtgaa 416
Db 355 GAGGACCTTGGAGAGGAGGAGTGTGACCTTGTGTGACTGGGAGATTCGAGAGGCGCT 296
Oy 417 ccgaggaacccgagcagcctctgagccgcgcgtacatagaacctgagctgaagtgagc 476
Db 295 CCGGAGAACCCGAGCAGCCTCTGCGCGCCGCTACATGAAACCTGAGCTCAGGAGCTTCC 236
Oy 477 cctccatcgtttgagaagcgtcctcagatgaatcctcactgctgaacctcctcctcagcc 536
Db 235 CTCTCATCGTTTGAGAACGCTCATGATGAATATCTCAACTGCTGAACTCTCTTCAAGCC 176
Oy 537 cactctgtgagcaagcttctagttagaagagctgtgataagtgatgaggaaggaact 596
Db 175 CACTCTGTGTGAGCAAGCTTCTAGTAGACACTCTTGGATAGTCAATGATGATGAGAGGAACT 116
Oy 597 gttagcaattgaagacagaaacccgagatgctg 628
Db 115 GTTGACAATGTGAAGACAAACCCGGGTAGTG 84

```

## RESULT 9

```

AA154039/c
ID AA154039 standard; DNA; 499 BP.
XX
AC AA154039;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #22725 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
WO200157272-AZ.
XX
09-AUG-2001.
XX
30-JAN-2001; 2001WO-US00663.
XX
04-FEB-2000; 2000US-0180312.
XX
26-MAY-2000; 2000US-0207456.
XX
30-JUN-2000; 2000US-0608408.
XX
03-AUG-2000; 2000US-0633366.
XX
21-SEP-2000; 2000US-0234687.
XX
27-SEP-2000; 2000US-0236359.
XX
04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX
XX analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID NO 22725; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).

```

CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX  
 SO Sequence 499 BP; 116 A; 142 C; 126 G; 115 T; 0 other;

Query Match 13.5%; Score 455.2; DB 22; Length 499;  
 Best Local Similarly 99.3%; Pred. No. 4.2e-103;  
 Matches 457; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Oy 169 atgtcgaatggatattccagagagaatctccgatactcactgctgctcagggc 228
Db 499 ATGTGCAATGGGATATTCACAGACAGAGAAATTCGGCTATCTCATCTGTGCTTCAAGGCG 440
Oy 229 aggttgaanaatgatactcagagtgagtgagctgtgtgtgactgaactgaccttctgctgca 288
Db 439 AGGCGAAATGTACATCCAGAGAGTGTGAGGCTGTGCTGAGTACCTGACCTTTCCTCCGCA 380
Oy 289 gaggltgaaggaagcagatccaagagacagtcgacacctccgaggaacatgcagcagltgaa 348
Db 379 GAGGTGAAGGAGCAGATTCAGAGGACAGTCCGCACTCCGGGAAACATGAGGAGTGGAA 320
Oy 349 ctgctgtgagcacccttgaggaagagagtcgagccttgatgagctgaggaattctg 408
Db 319 CTGCTGTGAGCACCCTTGAGAGAGGAGTGTGAGCACTGTGAGTGTGAGTGTGAGTGTG 260
Oy 409 gaggccctccgagaaacccgagcagcctctgagccgcgcgtacatagaacctgaagctcag 468
Db 259 GAGGCGCTCCGAGAACCCGAGCAGCCTCTGTGCGCCGCTACATGAACCTGAGACTACG 200
Oy 469 gaattgcccctcagatcgtttgagaagcgtcagatgaatatactcgaactgctgaacctc 528
Db 199 GACTTGCCCTCTCCATGCTTGTGAGAACGCTCATGATGAATATCTCAACTGCTGAACCTC 140
Oy 529 cttagcccaactctggtgagaaagcttctagttagaagagcttgagtaagtgatgagag 588
Db 139 CTTCAGCCCACTGTGTGTGAGCAACCTTCTAGTAGAACGCTTGTGATGAATGATGATGAGAG 80
Oy 589 gaggaaactgtgacaattgaagacagaaacccgagatgctg 628
Db 79 GAGGAGACTGTGACATTTGAAGACGAAACCCGGGTAGTG 40

```

## RESULT 10

```

AAA64601
ID AAA64601 standard; DNA; 458 BP.
XX
AC AAA64601;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of a human OLD-139 polynucleotide.
XX
XX OLD-35; OLD-64; OLD-137; OLD-139; OLD-142; OLD-175; cancerous phenotype;
XX
XX cellular senescence; terminal differentiation; growth suppression;
XX
XX aging process; type I interferon; cancer cell; tissue regeneration; ss.
XX
OS Homo sapiens.
XX
WO2000046391-A2.
XX
10-AUG-2000.
XX
02-FEB-2000; 2000WO-US02920.
XX
02-FEB-1999; 99US-0243277.
XX
(UYCO ) UNIV COLUMBIA NEW YORK.
XX
XX Fisher PB, Leszczynska M;
XX

```



Db 308 ccagaaattcaagaaactllacatlgtaagaanaaacaagaactgcaagaagtgtgc 367  
 QY 3024 cgactatcaataaatgttgaat 3047  
 Db 368 cgactatcaataaatgttgaat 391

## RESULT 12

AAV89625  
 ID AAV89625 standard: cDNA: 374 BP.

AAV89625:

DT 15-FEB-1999 (first entry)

DE EST clone C0522.

Human: secreted protein: expressed sequence tag; EST; haematopoiesis;  
 tissue growth; actinin; inhibin; chemotaxis; chemokinesis; haemostatic;  
 receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
 gene therapy; ss.

XX Homo sapiens.

XX MO9845436-A2.

XX 15-OCT-1998.

XX 10-APR-1998: 98MO-US06955.

XX 10-APR-1997: 97US-0838821.

XX (GENE) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Racie JA, Spaulding V, Treacy M;

XX WPI: 1999-070077/06.

XX New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries.

XX Claim 1: Page 277: 618pp: English.

XX The present sequence represents a human expressed sequence tag (EST).  
 CC The polynucleotide, which is a secreted EST, and the encoded protein  
 CC are predicted to have useful biological activities which would make  
 CC them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is  
 CC given. Suggested activities include nutritional activity, immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, actinin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 374 BP; 146 A; 47 C; 88 G; 93 T; 0 other;

Query Match 10.9%; Score 365.8; DB 20; Length 374;

Best Local Similarity 99.5%; Freq. No. 5.4e-81;

Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1824 gaactctagaataatgacaagaattcaactatgtaaatgagtcgaatgca 1883

DB 2 ggaactctagaataatgacaagaattcaactatgtaaatgagtcgaatgca 61

QY 1884 ttgtgaactcaaccctatgacaatgagccattcaaatggaanaaaagtcga 1943

DB 62 ttgtgaactcaaccctatgacaatgagccattcaaatggaanaaaagtcga 121

QY 1944 aggaatctgcaagaacgctgttgcgagaacatttgaaggaatgagccctaca 2003  
 Db 122 aggaatctgcaagaacgctgttgcgagaacatttgaaggaatgagccctaca 181  
 QY 2004 aattaatgacacacatttgcgaatgagatgctatcctacatccttgaacttataatga 2063  
 Db 182 aattaatgacacacatttgcgaatgagatgctatcctacatccttgaacttataatga 241  
 QY 2064 agagaagaataagaagcttgcgaatcatalagaagatgagatgagagtgatgata 2123  
 Db 242 agagaagaataagaagcttgcgaatcatalagaagatgagatgagagtgatgata 301  
 QY 2124 gttatgtagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 2183  
 Db 302 gttatgtagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 361  
 QY 2184 agatagact 2192  
 Db 362 agatagact 370

## RESULT 13

AAF66499  
 ID AAF66499 standard: cDNA: 387 BP.

XX AAF66499:

XX 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 2255.

XX Human: cytostatic; gene therapy; colon cancer; prostate cancer;  
 KW breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.

XX MO200102568-A2.

XX 11-JAN-2001.

XX 30-JUN-2000: 2000MO-US18374.

XX 02-JUL-1999: 99US-0142310.

XX 02-JUL-1999: 99US-0142311.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Williams IT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;  
 PI Reinhard C, Randozzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;  
 PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;  
 PI Kita D, Garcia V, Jones LW, Strache-Grain B;

XX WPI: 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a  
 PT mammalian cell and detecting cancer, particularly of the colon or  
 PT prostate, comprises 3351 human polynucleotide sequences -

XX Claim 9: Page 874: 1046pp: English.

XX The present sequence is one of 3351 sequences in a library of human  
 CC polynucleotides. The library is used to detect differentially expressed  
 CC genes correlated with a cancerous state of a mammalian cell and can  
 CC detect colon, prostate, breast and lung cancer. The library can be used  
 CC to produce probes for detection of mRNA and to produce additional copies  
 CC of the polynucleotides. The probes can be used for chromosome mapping of  
 CC the polynucleotide and for detection of transcription levels. Ribozymes  
 CC or antisense oligonucleotides can be generated. The polynucleotides and  
 CC their gene products are used as genetic or biochemical markers (e.g. in  
 CC blood or tissues) that will detect the earliest changes along the  
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
 CC preventive interventions. The polynucleotides, polypeptides and



PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI: 2001-488901/53.

XX Human genome derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX

PS Claim 25: SEQ ID No 9846; 487pp; English.

XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENP). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 292 BP: 108 A: 31 C: 71 G: 82 T: 0 other:

Query Match 8.4%; Score 281.4; DB 22; Length 292;  
Best Local Similarity 97.9%; Pred. No. 4.1e-60;  
Matches 285; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1934 ctgcaaaaaggaatcgcagaagaacgtgttgcagaacattgaggaagtacatg 1993  
DB 1 ctgcaaaaggaagaaatcgcaagaacgtgttgcagaacattgaggaagtacatg 60  
QY 1994 aggcctcaaatgaatgacacattggaatgagatgacgtatctcttgaactt 2053  
DB 61 aggcctcaaatgaatgacacattggaatgagatgacgtatctcttgaactt 120  
QY 2054 tctataatgaagaagaataaagaattgcagtcataagaatgatagtgaggggtg 2113  
DB 121 tctataatgaagaagaataaagaattgcagtcataagaatgatagtgaggggtg 180  
QY 2114 gtgatgatgagatgtgtgagtgatgaagatgagatgatttaagaacattgaac 2173  
DB 181 gtgatgatgagatgtgtgagtgatgaagatgagatgatttaagaacattgaac 240  
QY 2174 tggatgaacacagatgattctcactgacttattttgaaacataaaa 2224  
DB 241 tggatgaacacagatgattctcactgacttattttgtaagaattcaaa 291

Search completed: January 17, 2002, 03:26:40  
Job time: 7664 sec



...

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3365	100.0	3380	9	AF095844	AF095844 Homo sapi
2	3338.8	99.2	3373	9	AY017378	AY017378 Homo sapi
3	3194.4	35.5	1213	9	BC007966	BC007966 Homo sapi
4	1186	35.21	2050	10	BC004031	BC004031 Mus muscu
5	803	23.9	1258	6	AX098232	AX098232 Sequence
6	803	23.9	1270	6	AX098236	AX098236 Sequence
7	663	19.7	3652	6	AX098234	AX098234 Sequence
8	663	19.7	3704	6	AX098238	AX098238 Sequence
9	623.2	18.5	95417	9	AC010876	AC010876 Homo sapi
10	371.2	11.0	387	6	AX071769	AX071769 Sequence
11	356.8	10.6	392	6	AX071783	AX071783 Sequence
12	303	9.0	163681	9	AC007750	AC007750 Homo sapi
13	285.4	8.5	301	6	U50897	U50897 Sequence
14	269.2	8.0	377	6	A74554	A74554 Sequence
15	269.2	8.0	377	6	A77533	A77533 Sequence
16	223.4	6.6	2356	10	AF316999	AF316999 Mus muscu
17	178	5.3	2613	9	AF021416	AF021416 Homo sapi
18	148.6	4.4	3065	9	AF038963	AF038963 Homo sapi
19	141.8	4.2	160	6	A74468	A74468 Sequence
20	141.8	4.2	160	6	A77447	A77447 Sequence
21	130	3.9	112361	9	AC027316	AC027316 Homo sapi
22	130	3.9	143139	9	AC008413	AC008413 Homo sapi
23	124.6	3.7	193	6	U50896	U50896 Sequence
24	114	3.4	4024	4	AF181119	AF181119 Sus scrofa
25	109.6	3.3	263	9	HS080236	AL137608 Homo sapi
26	109.6	3.3	3026	9	AC023661	AL137608 Homo sapi
27	72.4	2.2	100635	9	AC003104	AC003104 Homo sapi
28	69.4	2.1	7218	6	U64494	U64494 Sequence
29	68	2.0	10480	10	AF317000	AF317000 Mus muscu
30	68	2.0	85440	2	AC034264	AC034264 Mus muscu
31	68	2.0	110000	2	AC021632_0	AC021632 Mus muscu
32	68	2.0	110000	2	AC021632_1	Continuation (2 of
33	68	2.0	170728	2	AC023804	AC023804 Mus muscu
34	68	2.0	183776	2	AL591469	AL591469 Mus muscu
35	68	2.0	263776	2	AC087335	AC087335 Mus muscu
36	67.4	2.0	38141	8	SPAC884	Z66569 S.pombe chr
37	67.4	2.0	41803	8	SPCC584	AL032824 S.pombe c
38	64.8	1.9	205429	2	AC005506	AC005506 plasmodiu
39	64.4	1.9	108908	3	IFMAL3P8	AF034560 plasmodiu
40	63.8	1.9	1815	3	AF322516	AF322516 plasmodiu
41	63	1.9	5815	8	AF187317	AF187317 Arabidops
42	63	1.9	6184	8	AF292940	AF292940 Arabidops
43	62.8	1.9	14339	3	AE001414	AE001414 plasmodiu
44	62	1.8	5852	9	AB023145	AB023145 Homo sapi
45	62	1.8	7037	9	AB028449	AB028449 Homo sapi

## RESULT

LOCUS	AF095844	3360 bp	mRNA	PR1	26-NOV-2000
DEFINITION	Homo sapiens melanoma differentiation associated protein-5 (MDA5-mRNA, complete cds.				
ACCESSION	AF095844				
VERSION	AF095844.1	G1:11344593			

**SOURCE**

ORGANISM      Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.

## REFERENCE

**AUTHORS**

LIFE

JOURNAL

## REFERENCE

## AUTHORS

TITLE

**JOURNAL:** Submitted (10-SEP-1998) Urology, Columbia University, P&S, 630  
W168th, New York, NY 10032, USA  
**FEATURES** Location/Qualifiers  
**SOURCE** 1. 3380







QY	1515	caagaagcagatgataataaactatagggcattatttgatgcgaagtttgaaaaaaca	1514
Db	1501	CAAGAAGCAGATGATATAATCAATCATGAGGCATTAATTGATGCGAAGTTGAAAAACA	1500
QY	1575	tagactaaagaagaacaacacagtgatccccctctcagataacggactaaagc	1634
Db	1561	TAGACTCAAGAAACAAAACACAGTGATTCCTCTCGATATCTGGGCTTACAGC	1620
QY	1635	ctcaactggtgttgaggggcccagaagccaagctggaagacacattttaaact	1694
Db	1621	CTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACT	1680
QY	1695	atgtgcgaattcttgatgacttaacttaaacctgttlaagaaaccttgatcaactgaa	1754
Db	1681	ATGTGCCAATCTTGATGACTTATTAATCTTTAAAGAAAACCTTGATCAACTGAA	1740
QY	1755	aaaccaatatacgaagccatgcgaagaagtttgcattgagaatgaacagaagaatcc	1814
Db	1741	AAACCAATATACGAGGCCATGCGAAGAAATTGGCATTTGAGATGCCAACCGAAGATCC	1800
QY	1815	atttaagaagaaactctagaataatgacaagatccaactatgttcaaatgattcc	1874
Db	1801	ATTTAAGAAGAACTCTAGAATAATGACAAAGATTCAACTATATGTGCAATGAGTCC	1860
QY	1875	aatgtcgaattcttggaactcaaccctatgacaacaatggtccattcaatggaaaaaagc	1934
Db	1861	AATGTCAATTTCTGGAACCTACCCCTATGACCATGGCCATTCAAATGTGAAAAAAGC	1920
QY	1935	tgcanaaaagaaatcgcgaagaagctgtttgtgcagaactttgagaagaatcaatga	1994
Db	1921	TGCAAAAAGAGAAATGCGAACAAAGCTGTTTGTGAGAACATTGTAGCAACTACATATA	1980
QY	1995	ggccctacaattaatgacacaacttcgaatgatagtgcatactcattgaaactt	2054
Db	1981	GGCCCTACAATTAATGACACCAATTCGAATGATAGATGCGTATACGACTTGGAAATTT	2040
QY	2055	ctataatgaaagaagaataagaagtttgcagtcataagaagatgatatgataaggtgtg	2114
Db	2041	CTATATGTAAGAGAAATATGAAGATTTGCAGTCATAGAAATGATATGATGACGGTGG	2100
QY	2115	tgatgatagatcattgatatgatatgataagaatgagatgatttlaagaacacttgaact	2174
Db	2101	TGATGATGAGTATTTGTATGCTGCTCATGACATGCGATTTTAAAGAAACCTTTGAAACT	2160
QY	2175	ggaatgaacacatagatcttctcagacttattcttgaacaacaataatgttgaagaag	2234
Db	2161	GGATGAACACAGATGATGATTTCTCATGACTTATTTTGAAGAACATATAATGTGTAAAG	2220
QY	2235	gctgtcgtgaaaaaaccagaataatgaataagctgcaccaataaagaataacataat	2294
Db	2221	GCTGTGCTGTAACCCACAGATATGAATAATGAAGAGCTGACCAATATTAAGAAATGCCATAT	2280
QY	2295	ggaagaacataactagagcttagaatcagaagaaggaataaactttacaaaaacagaca	2354
Db	2281	GGAAGCAATATCTAGGAGCTGAGCAATGAGCAGCAGGAATTAATCTTTACAAAACACGACA	2340
QY	2355	gagtgacataatgcgtctccagtggaattacttgaataatgaataatgtcgtgaatgaggt	2414
Db	2341	GAGTGCAATATGCGCTTTCGCCAGTGATTACTGAAATGAAAAATTTGCGTGAAGTAGAGT	2400
QY	2415	caaaagccacacatctgattgagctgtgacacagcagctgattccaaccaatgacagaa	2474
Db	2401	CAAAAGCCACACATCTGATTTGAGCTGAGCACACACAGTGAATTCMAAACCATATCACAGAA	2460
QY	2475	tgaacaaaaaagatcttgatgaatttgcattgcattggaataatcaactctgattatgcac	2534
Db	2461	TGAACAAAAAAGTCAATTATTAATTTCCGCTGGAATAATAATCTGCTTATGCTATAC	2520
QY	2535	caacgttgcagaagaagctcgtgatatlaagaatgatacatgatatccgtatgctt	2594
Db	2521	CACAGTGCAGACAGAGTGTCGATATTTAAGAATGTAACTATGTTATTCGTTATATGCTCT	2580
QY	2595	cgctcaacaatgaatagccatggtccagcccggtgtcgagcccaagctgatatgagagcac	2654

Db	2651	CGTCAACCAATGAAATATAGCCATGCTGACAGGCCCGCTGGTCGACCCAGAGACTGATGAGACAC	2640
QY	2655	ctacgtccctggttctcaccagttggttcaaggatattcgaaatcgaaatgaagttaatgatt	2714
Db	2661	CTACGTCCTCGGTTGCTGCACACTGGTTTCAGGAGGTTATGCAAGCTGACAGACTATATGATT	2700
QY	2715	ccgagagaagaatgatgataaagctatacatctggttcacaataatgaaccagagagta	2774
Db	2701	CCGAGAGAGAGATGATGATATAAAGCTATACATTGGTTCCAAATAATGAAACAGAGGAGTA	2760
QY	2775	tggtcattaaagtttttggaaattacagatgccaagatataatgysaaaagaataatgaaccac	2834
Db	2761	TGCTCATAAAGATTTTGGAAATTTACAGATGCAAAAGATATATGAGAAAAGAAATGAAACCAA	2820
QY	2835	gagaataattccaaagcattacaagaataaccataaccataacttctcttgcacaaa	2894
Db	2821	GAGAAATTTGTCGCAAGCATTTCAAGAAATACCCATCATTATTAACCTTCTTGCCAAAAA	2880
QY	2895	ctcgagtgctgaagcctgtgttcctggggaagatataccatgtaattgaagaaaatgcatcagt	2954
Db	2881	CTGCACTGTGCTAGCCCTGTTCTGGGGAAAGATATCCATGTAATTTGAGAAAATGCATCCGT	2940
QY	2955	caatatagcccccagaattccaagaaacttcaatttgaatgaagaaaacaaagcactgcacaa	3014
Db	2941	CAATATGACCCCGAATTTCAAGGAACCTTTACATTGTAAAGAGAAAACCAACACCTGCACAAA	3000
QY	3015	gaagtgtccgaactatacaataatgtgaaatcatctgcacatgtgtgccaggtcttgag	3074
Db	3001	GAACTGTCCGCACTATCAAAATTAATGTGTGAATATATCTGCACAAATGTGCGCCAGCTTGCGG	3060
QY	3075	aacaatgatgtgcacaaaggcttagatttgctcctgtgtccaaataagaaatgttgagt	3134
Db	3061	AACATGATGCTGTCACAAAGCGTTAGATTTCCTTGCTCCAAATTAAGAAATTTGTAGT	3120
QY	3135	ggtttcaaaaataattcaacaagaanaacatacaaaaagtggttagaattactatcac	3194
Db	3121	GGTTTCAAAAATATTTCAACAAAGAAACATATCAAAAAAGCTGGGTACAAATTAACCTATCAC	3180
QY	3195	atttcccaatcttgactatcagaatgtgttatttgaatgatgaattagacacttgat	3254
Db	3181	ATTTCGCAATCTTGCACTATTCAGAAATGCTGTTATTATTAGTATGACAGATTAGCACTTGAT	3240
QY	3255	tgaagatctctttaaataactatcagttaaacattlaatatgattatgattatgattc	3314
Db	3241	TGAAGATCTCTTAAATACTATCATGTTAAACATTTAATATGATTAGATTAAATGATATTC	3300
QY	3315	attatgctacagaactgacataaagaatacaataaataatgattcttactct	3364
Db	3301	ATTATGTCACAGAACTGCATATAGCAATCAATAAATAATGATTGTTTACTCT	3350

RESULT 3

BC007966 1213 bp mRNA PRI 12-JUL-2001

LOCUS Homo sapiens, Clone IMAGE:4304805, mRNA.

DEFINITION BC007966

ACCESSION BC007966.1 GI:14044090

VERSION

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1213)

AUTHORS Strausberg,R

TITLE Direct Submission

JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapops-remail.nih.gov](mailto:cgapops-remail.nih.gov)

Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland.  
 Web site: <http://www.nisc.nih.gov/>  
 nisc\_mgenhgr1.nih.gov

Contact: [nisc\\_mgenhgr1.nih.gov](mailto:nisc_mgenhgr1.nih.gov)  
 Shevchenko, V., Wetherby, K.D., Bouckstrom, Sternberg, S.M.,  
 Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dierlich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stancirpop, S., Thomas, P.J.,  
 Tlonson, R.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IMAGE Plate: 20 Row: P Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

## SOURCE

location/Qualifiers  
 1..1213  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4304805"  
 /tissue\_type="uterus, leiomyosarcoma"  
 /clone\_lib="NIH\_MGC\_46"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOB7"  
 BASE COUNT 386 a 247 c 285 g 295 t  
 ORIGIN

Query Match 35.5% Score 1194.4; DB 9; length 1213;  
 Best Local Similarity 99.5%; Pred. No. 3e-242;  
 Matches 1198; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 380 ggcaccttggatggaactcgaatcgtggaagccctcgaagaacgcgcgcctctgg 439
    |||||||
DB 9 ggcaccttggatggaactcgaatcgtggaagccctcgaagaacgcgcgcctctgg 68
QY 440 ccgcgcgtacatgaacccctgagctcagcgcctcgcctcgcctcgtttgaaagcgtc 499
    |||||||
DB 69 ccgcgcgtacatgaacccctgagctcagcgcctcgcctcgcctcgtttgaaagcgtc 128
QY 500 atgataatctcgaactcgtcgaacccctcgcctcgcctcgcctcgtttgaaagcgtc 559
    |||||||
DB 129 atgataatctcgaactcgtcgaacccctcgcctcgcctcgcctcgtttgaaagcgtc 188
QY 560 ttgagacgctcttgataagtcagtcaggaaggaactgtgacattgaagacagaaccc 619
    |||||||
DB 189 ttgagacgctcttgataagtcagtcaggaaggaactgtgacattgaagacagaaccc 248
QY 620 ggaattcgtcgtcgaagaacaaatgaatgaatcagtgatgaagagagctactaaaga 679
    |||||||
DB 249 ggaattcgtcgtcgaagaacaaatgaatgaatcagtgatgaagagagctactaaaga 308
QY 680 ttgtcagaanaaagaacggtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 739
    |||||||
DB 309 ttgtcagaanaaagaacggtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 368
QY 740 atgaactgtcctcaagaatgaacgctctgattgtcagaagaacatgacagatgaga 799
    |||||||
DB 369 atgaactgtcctcaagaatgaacgctctgattgtcagaagaacatgacagatgaga 428
QY 800 attatccaaagtgatggtcctcagaatgaagaagcaactcttcaaccacagctcagc 859
    |||||||
DB 429 attatccaaagtgatggtcctcagaatgaagaagcaactcttcaaccacagctcagc 488
QY 860 caaatctgagagagagatctgaggaatgaggaataactcaatcaatcatttggcag 919
    |||||||

```

```

DB 489 CAATCTGGAGAGAGAGCTGTGGGGCATGAGAAATTAACATCATGAAATCTTTGGCAG 548
QY 920 attctcgtatgcttcaagaatcagacacaaattgagcagaagagatgacgtcttag 979
    |||||||
DB 549 ATTCTTGCTGATGTTTCAATATCAGACACAACTTTGGAGAAAGAGTGTAGCTGCTTAG 608
QY 980 atgaagctcttggacataaacagacacatgagcagatgattcaggtaccatgggaatgatt 1039
    |||||||
DB 609 ATGAAGAGCTTGGACATTAACAGCAACATGGCGACGTGATTCAGCCATATGGAAGTGAAT 668
QY 1040 cagatgagaagaatgagtgagcagaagacatcccgagagccagacacatcagctcagctt 1099
    |||||||
DB 669 CAGATGAAGAGAAATGTGACAGCAAGACATCCCGAGCGCAGAGACTCTAGTACGCTT 728
QY 1100 accaaatgagaagttgcccagcagccttggaaggaagaatatacatatctgctcccta 1159
    |||||||
DB 729 ACCAAATGGAAGTGGCCAGCAGCAGCTTGGAGAGGAAGAAATATATATCTGCTCCCTTA 788
QY 1160 caaggaatggaagaaacccagagtggtcttaccattgcccagagatcattagacaagaaga 1219
    |||||||
DB 789 CAGGAGTGGCAAAACAGAGAGTGTGTTACATTTGCCAAGAGATCATTAGACAAAGAGA 848
QY 1220 aaaaagatctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1279
    |||||||
DB 849 AAAAAGCATCTGACCTGGAGAAAGTTATAGTCTTGCAATTAAGTACTGCTAGTTAAC 908
QY 1280 agctcttcgcagaagagttccacacattttagaagaatgatalcgtgtatagattaa 1339
    |||||||
DB 909 AGCTCTTCGCAAGAGAGTGTCAACCATTTTGAAGAAATGTATCTGTTATTGGATTAA 968
QY 1340 gtggtatgcccaacacgaataatcattcctcgaagtgatgacgtcgtcgtatattatla 1399
    |||||||
DB 969 GTGCTGATACCCCACTGAATAATATCATTTCCAGAAAGTTGTCAGATGCTGATATTATTA 1028
QY 1400 tcaatcagacgtcgaatcttcttgaagaactccctcttcaacttggaaaatgggaagatgctg 1459
    |||||||
DB 1029 TCAGTACAGCTCAATATCTTGAAACCTCCCTTTAACTTGGAAGAGAGAAAGATGCTG 1088
QY 1460 gtgtcaatgtcgaactcttccctcattcatttgaatgaatgcatcacacacaaag 1519
    |||||||
DB 1089 GTGTCAATGTGACAGCTTTCCCTCATTTATGATGATGATGATGATGATGATGATGATG 1148
QY 1520 aagcagtgataataacatcagatgagcattatgagcagaagttgaaacaaatagac 1579
    |||||||
DB 1149 AAGCAGTGATTAATTAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1208
QY 1580 tcaa 1583
    ||
DB 1209 AAAA 1212

```

## RESULT 4

BC004031 2050 bp mRNA R0D 12-JUL-2001

LOCUS BC004031 Mus musculus, clone IMAGE:3495361, mRNA, partial cds.

DEFINITION BC004031

ACCESSION BC004031.1 GI:13278455

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REMARK

COMMENT

house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2050)  
 Strausberg, R.  
 Direct Submission  
 Submitted (28-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-f@mail.nih.gov](mailto:cgaps-f@mail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.



Db	257	--CAGAGGTGAAGGAGCGAGATTCAAGAGGACAAATCCACCACCTCCGGGAGACATGACAGCGAGT	314
Qy	345	tgaactgctgcttgagcacccttggaaagaagagatctcggtaacttgattggatccggaatt	404
Db	315	tgaactgctgctgagacaccttggaagaagagatctcggtaacttgattggatccggaatt	374
Qy	405	cgttgaagccctccggagaacccggcagccctctgacgcgcgtacatcgaaccccgact	464
Db	375	cgtgaagccctccggagaacccggcagccctctgacgcgcgtacatcgaaccccgact	434
Qy	465	caaggacttgcctctccatcggttggaaagctcagaagaataatctccaaactgctgaa	524
Db	435	caagcacttgcctctccatcggttggaaagctcagaagaataatctccaaactgctgaa	494
Qy	525	ccctctcagcccaactcgttggacaagctcttgaatgaagaagctcttgaatgaatgcat	584
Db	495	cctctctcagcccaactcgttggacaagctcttgaatgaagaagctcttgaatgaatgcat	554
Qy	585	qgaagagaagactgttgcacaattgaaagacaaaacccgaattgctgctcagaaaaaaactg	644
Db	555	ggagagagaagactgttgcacaattgaaagacaaaacccgaattgctgctcagaaaaaaactg	614
Qy	645	aaatgaatcaagctgaagaagagctactaaaagaagatttggagaagaagaagctggtctc	704
Db	615	aaatgaatcaagctgaagaagagctactaaaagaagatttggagaagaagaagctggtctc	674
Qy	705	tgcatttctgaatgttctctgcacaacagagaagaacatgaactgttccaaagattaaacg	764
Db	675	tgcatttctgaatgttctctgcacaacagagaagaacatgaactgttccaaagattaaacg	734
Qy	765	ctctgaattgtcagaagaagcaatgcagagattgagaattatcacaaagtltgatgtctca	824
Db	735	ctctgaattgtcagaagaagcaatgcagagattgagaattatcacaaagtltgatgtctca	794
Qy	825	agtgagaagaagcaactctcttcaaccacagcttcacgccaactctgagaagaagcttgcg	884
Db	795	agtgagaagaagcaactctcttcaaccacagcttcacgccaactctgagaagaagcttgcg	854
Qy	885	catgagaataaactcaatcaagaatcaatcttlttgacagatctctctgatttcaag	937
Db	855	catgagaataaactcaatcaagaatcaatcttlttgacagatctctctgatttcaag	907

RESULT 6

AX098236 LOCUS AX098236 1270 bp DNA PAT 30-MAR-2001

DEFINITION Sequence 11 from Patent WO0118208.

AX098236 ACCESSION

AX098236.1 GI:13515371

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Peyman,J.A., da Silva,A., Hochman,P. and Hsu,A.

JOURNAL Interferon induced polynucleotides and proteins encoded thereby

TITLE Patent: WO 0118208-A 11 15-MAR-2001;

CURATOR Curagen Corporation (US); BIOGEN, INC. (US)

FEATURES

source

1..1270

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1..927

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC35442.1"

/db\_xref="gi:13515372"

/translation="MPAAVNSNTPRQVYIKSPVNSGGRPCSKMAADSGTHISPV

POTTRPAMENPLPILSKRIYENVYFHHREPLISLVIQGGENVHPEYKEQIORTVAT

SGNQAQVELLLSTLEKGVNHLGWTFVEALRKFGSLAARVNNPELTDLSPSENA

HDEVLAILLLOPTLVKRLVRIYLDKCMREELITLDNRILAAENNGNSGVRELL

KRIYQKNMFSAPLIVNLRQTNNEILVDELFGSDKFSNAILENISYVYIDPVEQLILS









Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
 Summary Statistics  
 Center project name: H\_NH0214A04

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tano, M., Catanesi, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-576116, 200 bp overlap; the clone sequenced to the right is RP11-204024. Actual start of this clone is at base position 126545 of RP11-576116; actual end is at base position 81443 of RP3-404K21.

The sequence RP11-214A4 from base position 23333 to 23415 was derived from PCR off clone DNA.

#### FEATURES

source  
 Location/Qualifiers  
 1..95417  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-214A4"  
 /clone\_lib="RPCI-11"  
 248..269  
 /rpt\_family="AT\_rich"  
 358..383  
 /rpt\_family="AT\_rich"  
 695..726  
 /rpt\_family="AT\_rich"  
 740..950  
 /note="similar to Mus musculus EST AA123960 (NID:g1682517)  
 mp78b04.r1"  
 763..950  
 /note="similar to Mus musculus EST AA656698 (NID:g2592852)  
 vp95c09.r1"  
 2364..2505  
 /rpt\_family="MERL\_type"  
 2714..2828  
 /rpt\_family="MIR"  
 2916..3111  
 /rpt\_family="MIR"  
 5265..5370  
 /rpt\_family="MIR"  
 6509..6734  
 misc\_feature

misc\_feature  
 mp78b04.r1  
 6509..6710  
 /note="similar to Mus musculus EST AA656698 (NID:g2592852)  
 vp95c09.r1"  
 7292..7336  
 /rpt\_family="(CA)n"  
 7585..7664  
 /rpt\_family="AT\_rich"  
 8919..8968  
 /rpt\_family="L2"  
 10445..10499  
 /rpt\_family="L2"  
 11237..11409  
 /rpt\_family="L1"  
 11552..11585  
 /rpt\_family="AT\_rich"  
 11938..12042  
 /note="similar to Mus musculus EST W29262 (NID:g1309410)  
 mc22c08.r1"  
 12057..12081  
 /rpt\_family="AT\_rich"  
 13116..13146  
 /rpt\_family="AT\_rich"  
 13371..13454  
 /rpt\_family="L1"  
 13455..13787  
 /rpt\_family="L1"  
 14714..14943  
 /rpt\_family="MIR"  
 15410..15462  
 /rpt\_family="AT\_rich"  
 17142..17196  
 /rpt\_family="(CA)n"  
 17798..17989  
 /rpt\_family="MERL\_type"  
 18156..18391  
 /rpt\_family="L2"  
 20698..20741  
 /rpt\_family="AT\_rich"  
 21490..21789  
 /rpt\_family="Alu"  
 21846..21954  
 /rpt\_family="L2"  
 21957..22046  
 /rpt\_family="GA-rich"  
 22125..22147  
 /rpt\_family="AT\_rich"  
 22566..22843  
 /rpt\_family="Alu"  
 23662..23943  
 /rpt\_family="Alu"  
 23946..23988  
 /rpt\_family="(TAAAA)n"  
 24002..24411  
 /rpt\_family="MERL\_type"  
 24948..24971  
 /rpt\_family="AT\_rich"  
 25083..25227  
 /note="similar to Mus musculus EST W29262 (NID:g1309410)  
 mc22c08.r1"  
 26227..26432  
 /rpt\_family="MIR"  
 26469..26755  
 /rpt\_family="MALR"  
 26756..27051  
 /rpt\_family="Alu"  
 27052..27155  
 /rpt\_family="MALR"  
 27157..27613  
 /rpt\_family="MALR"  
 27897..28473  
 /rpt\_family="MALR"



source 1.387  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 143 a 51 c 91 g 102 t  
ORIGIN

Query Match 10.6% Score 356.8; DB 6; Length 387;  
Best Local Similarity 96.8%; Pred. No. 3.4e-65;  
Matches 364; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1882 gatcttggaactcaaccatgaacatgagccatcaatgagaaataaactgcaaa 1941  
DB 11 GATTGTGCACTGACACCTATGACCAATGGCCATTCAAATGGAATAAAGTSCAAAA 70

QY 1942 aaaggaatcgcaagaacgigtgttcagaaacattgaggaagtacatgagcccta 2001  
DB 71 GAAGGAATCGCAGAAACGTGTTGTCAGAACATTTGAGGAAGTACCAATGAGCCCTA 130

QY 2002 caaatatgacacattcgatgataatgatactcattgaaactttcattat 2061  
DB 131 CAAATTTATGACACATTCGATGATGATGATGATGATGATGATGATGATGATGAT 190

QY 2062 gaaggaagaataaagaatttcagatcagatagaagaatgataatgataatgataat 2121  
DB 191 GAAGGAAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250

QY 2122 gacttttgatggtgatagaagaatgataatgataatgataatgataatgataat 2181  
DB 251 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310

QY 2182 acagataatctcattgacttatttttgaacataaataatgtaaaagctgct 2241  
DB 311 ACAGTATCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 370

QY 2242 gaaacccagaatg 2257  
DB 371 GAAACCCAGATATG 386

RESULT 12  
AC007750 163681 bp DNA PRI 02-OCT-2000  
LOCUS Homo sapiens BAC clone RP11-576116 from 2, complete sequence.  
AC007750  
AC007750.3 GI:6094634  
VERSION  
KEYWORDS  
HTG.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
99063792  
2 (bases 1 to 163681)  
Cotton, M., Maupin, R., Hawkins, M. and Harkins, R.  
The sequence of Homo sapiens BAC clone RP11-576116  
unpublished  
3 (bases 1 to 163681)  
Waterston, R.H.  
Direct Submission  
Submitted (05-JUN-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 163681)  
Waterston, R.H.  
Direct Submission  
Submitted (22-OCT-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
5 (bases 1 to 163681)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Waterston, R.  
Direct Submission  
Submitted (02-OCT-2000) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Oct 22, 1999 this sequence version replaced 91:510389.  
-----  
Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: saplens@wustl.wustl.edu  
Summary Statistics  
Center project name: H\_NH0576116  
-----

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D.  
McPherson, Department of Genetics, Washington University, St. Louis  
MO. For additional information about the map position of this  
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
The RPC1-11 human BAC library was made from the blood of one male  
donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,  
Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved  
approach for construction of bacterial artificial chromosome  
libraries. Genomics 51:1-8. The clones may be obtained either from  
Research Genetics, Inc. (http://www.rgsen.com) or Pletier de Jong  
and coworkers at the Roswell Park Cancer Institute  
(http://bacpac.med.buffalo.edu)  
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-178A14. Actual start of  
this clone is at base position 1 of RP11-576116; actual end is at  
base position 163681 of RP11-576116.

FEATURES  
SOURCE  
1.163681  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-576116"  
/clone\_1bp="RPC1-11"  
216..322  
/rpt\_family="MER81"  
1271..1298  
/rpt\_family="(CAGCT)n"  
1536..1694  
/rpt\_family="MIR"  
2098..3518  
/rpt\_family="12"  
3614..3902  
/rpt\_family="Alu"  
4090..4568  
/rpt\_family="MER1\_type"  
5128..5148  
/rpt\_family="AT-rich"  
5263..5299  
/rpt\_family="(CA)n"  
5417..5648  
/rpt\_family="LI"

repeat_region	7093. .7556	/rpt_family-"L2"	
repeat_region	8702. .9187	/rpt_family-"MalR"	
repeat_region	9386. .9586	/rpt_family-"MLR"	
repeat_region	10580. .10739	/rpt_family-"L2"	
repeat_region	10792. .11064	/rpt_family-"Alu"	
repeat_region	11136. .12912	/rpt_family-"L1"	
msc_feature	12894. .12899	/note="match to EST A1077564 (NID:g3411972) 0233q05.x1"	
repeat_region	12982. .13011	/rpt_family-"(TTG)n"	
repeat_region	13109. .13133	/rpt_family-"(TTA)n"	
repeat_region	13134. .13419	/rpt_family-"Alu"	
repeat_region	13833. .14130	/rpt_family-"Alu"	
repeat_region	14437. .14573	/rpt_family-"MER1_type"	
repeat_region	14742. .14887	/rpt_family-"L1"	
repeat_region	15098. .15234	/rpt_family-"L1"	
msc_feature	15444. .15460	/note="match to EST A1922427 (NID:g5658391) w006b10.x1"	
repeat_region	15449. .15475	/rpt_family-"(T)n"	
repeat_region	15476. .15695	/rpt_family-"L1"	
repeat_region	15713. .15993	/rpt_family-"Alu"	
repeat_region	15994. .16013	/rpt_family-"(TAA)n"	
repeat_region	16039. .16131	/rpt_family-"L1"	
repeat_region	16140. .16281	/rpt_family-"L1"	
repeat_region	16284. .16327	/rpt_family-"(TG)n"	
repeat_region	16328. .16394	/rpt_family-"(CAT)n"	
repeat_region	16396. .16576	/rpt_family-"L1"	
repeat_region	16625. .16997	/rpt_family-"L1"	
repeat_region	17020. .17274	/rpt_family-"L1"	
repeat_region	17270. .17603	/rpt_family-"L1"	
repeat_region	17644. .18442	/rpt_family-"L1"	
repeat_region	18444. .20187	/rpt_family-"L1"	
repeat_region	20260. .20562	/rpt_family-"MER2_type"	
repeat_region	20563. .20850	/rpt_family-"Alu"	
repeat_region	20851. .20910	/rpt_family-"MER2_type"	
repeat_region	21194. .21218	/rpt_family-"(T)n"	
msc_feature	21204. .21219	/note="match to EST A1589439 (NID:g4598487) lr76a08.x1"	
msc_feature	21204. .21219	/note="match to EST A1582416 (NID:g4568313) lr97d11.x1"	
msc_feature	21204. .21219	/note="match to EST A1865901 (NID:g5530008) wk88b06.x1"	

Query Match	Best local Similarity	Matches 303: Conservative	9.0%: Score 303; DB: 9; Length 163681; Pred. No. 1,2e-53; Mismatches 0; Indels 0; Gaps 0;
repeat_region	/note="match to EST A1563407 (NID:g4569304) ts0d001.x1"	21618. .21730	
repeat_region	/rpt_family="MERL_type"	21800. .21840	
misc_feature	/rpt_family="AT-rich"	21991. .22005	
repeat_region	/note="match to EST A1682812 (NID:g4892994) wc66f03.x1"	21991. .22290	
repeat_region	/rpt_family="Alu"	22454. .22658	
repeat_region	/rpt_family="MALR"	22756. .23030	
repeat_region	/rpt_family="MALR"	23343. .23397	
repeat_region	/rpt_family="(CA)n"	23919. .23983	
repeat_region	/rpt_family="MIR"	24444. .24581	
misc_feature	/rpt_family="MERL_type"	24726. .24983	
misc_feature	/note="match to EST T10929 (NID:g391083)"	24730. .25234	
misc_feature	/note="match to EST A1682793 (NID:g4892975) wc66d05.x1"		
Query Match	9.0%: Score 303; DB: 9; Length 163681;		
Best local Similarity	100.0%; Pred. No. 1,2e-53; Mismatches 0; Indels 0; Gaps 0;		
Matches 303: Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY 3063	ccagccttggggaacaaatgatgtgcacaaagcttagatgtgctgtctcacaataag 3122		
Db 149238	CCAGCGTGGGGAACAAATGATGCTGCACAAAGGCTTAGATTGGCTTGCTCAAAATAG 149179		
QY 3123	gaatttlytagtlygtttcraaaaatattcaacaaagaacacatacaaaaatggttaga 3182		
Db 149178	GAATTTTGAGTGTTCACAAAAATATTCACAAAGAACAAATACAAAGAGTGGTAGA 149119		
QY 3183	attaccatacaatttcccaattcttgactatccaagatgcgttattatgagatgagga 3242		
Db 149118	ATTACTTACACATTTCCCAACTTGACATTCACAAATGCTCTTTATTAGCATGAGCA 149059		
QY 3243	ttagcaacttgatgaagatctctttaaatactatcaqttaaacaattlaattgattag 3302		
Db 149058	TTAGCAGCTTGATGAGAGATCTTTTAAATACTATCAGCTTAACATTTAATGATTAG 148999		
QY 3303	attaatgatcattatagtcacagaacctgacataaagaatcaataaaatgattgttact 3362		
Db 148998	ATTATGATTCATTATGCTACAGCAAGTCACATTAAGCAATCAATTAATGATTGTACT 148939		
QY 3363	ctg 3365		
Db 148938	CTG 148936		
RESULT 13			
LOCUS 150897	150897 301 bp DNA PAT 07-OCT-1997		
DEFINITION	Sequence 7 from patent US 5643761.		
ACCESSION	150897		
VERSION	150897.1 GI:2472600		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 301)		
TITLE	Fisher,P.B. and Jlang,H.		
JOURNAL	Method for generating a subtracted cDNA library and uses of the		
FEATURES	Patent: US 5643761-A 7 01-JUL-1997;		
source	location/Qualifiers		
	1..301		
	/organism="unknown"		
BASE COUNT	110 a 61 c 56 g 74 t		
ORIGIN			











FEATURES	source
LOCUS	827 bp mRNA
DEFINITION	602631817F1 NCI_CGAP_skn3 Homo sapiens CDNA clone IMAGE:4776869 5',
ACCESSION	U00000.1
VERSION	1.0
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (bases 1 to 827)
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strusberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: James Cleaver, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLNL0630 row: b column: 06 High quality sequence stop: 763. Location/Qualifiers 1..827

```

/lab.host="DH10B (T1 phage-resistant)"
/!note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5Kb. Library constructed by Life
Technologies. Note: this is a NCI-CGAP library."
BASE COUNT      254 a      186 c      204 g      183 t
ORIGIN
Query Match      19.4%: Score 652.8; DB 11; Length 827;
Best Local Similarity 96.1%: Pred. No. 8,2e-124;
Matches 798: Conservative 0; Mismatches 17; Indels 15; Gaps 12.
QY      426  cagcagccctctgcgcgcgcgtacacgaacctgaagctcaagagacttgcctctccac 485
          |||||||
Db      1    CGGACACCCCTTGCGCG-CCGCTACATGAAACCCCTAGCTCAGAGGATTTCCCTCTGCAC 59
QY      486  gtttgaagaagctcatalgaatataatcccaactgctgaacctctcttcag-cccactctg 544
          |||||||

```

Db	60	GTTCGAGACAGCCCTATCATGATGATATCTCCAACTGCTGAACCTCTTCAAGTCCCACTCTGG	119
Qy	545	tgagcaagctctcgtcttagagaagctcttgataagtgatcagtgaggaagcaactgtgacaa	604
Db	120	TGACAAACCTCTCTAGTTAGAGACGCTTGGATTAAGTCATGAGAGAGACCTTTGCAAA	179
Qy	605	ttgaagagaagaaccggatgtcgtctgcagaaaacaaatgaaatgaatcagtggtlaagag	664
Db	180	TTGAAGACAGAAACCGCATTTGCTCTCCAGAAAACAATGTGAAATGAATCAGGTGTAAGAG	239
Qy	665	agctactaaaagatgtgtgcagaaagaagaactgtgtctcgtacattctgaatgtcttc	724
Db	240	AGCTACTAAAAGCATTTGTGTCAGAAAGAAACTGTTCTCTGCATTTCTCAATGTTCTTC	299
Qy	725	gtcaaacacaggaacaatgaaactgtgtccaaagagtaaacagctctgattg--ctcagaag	782
Db	300	GTCMAACAGGAACCATGATGACTTGTCCACGAGTTAACAGGCTCTGATTGGCTCAGAAAAG	359
Qy	783	caatgcagagatgtagaattatcaaaagttgatgtgtcctcaagtgtgaagagcaactct	842
Db	360	CAATGCAAGAGATTAGAAATTTATCACAAAGTTGATGTGCTCAAGTGAAGACAACTTCT	419
Qy	843	ttcaaccacactttcagccaactctggga--gaaagaggtctggggatgtgaagaaataactcat	901
Db	420	TTCAACCAACAGTTTACGCCAAATCTGGACGGAAGGAGGTCTGGGGCATGTGACAAATTAACCTAT	479
Qy	902	cagaatcatctcttgca--gattctctgttaattcaagaatcagacacaagtcttgca--g	958
Db	480	CAGAAATCATCTTTTGGACAGATTCTTCTGTATTTCACAAATTCAGACAAAGTTTGGCAGC	539
Qy	959	aaggaagtgtcagctgtcttagatgaagactcttgca--cataacagacaatatgggcagtgat	1017
Db	540	AAGGAGGTGTAGCTGTCTTAGATGAAGTCTTGGACCATTAACACCAACATGGGCGAGTAT	599
Qy	1018	tcaaggcaccatgggaagaagtgtatgatgaa--gagaagtgtgcagcaagaatctcccgga	1076
Db	600	TCAGGCACCATGAGGGAAGTGAATTCAGATGAAGAAATGTGGCACCAAGACATCTCCCGGA	659
Qy	1077	gccaaactccagctcgaagctcttaaccaatlg--aaqtlgcacagcagccttgaaag9	1134
Db	660	GCCAAACTCCAGCTCAGGCTTACCAATGTGGCACAGTGTGCCAGCCAG--CTTGGCAGGG	718
Qy	1135	aagaatatcatcatgtgtcctccctacagggagtggaanaacagagatgtgctgtttacat	1194
Db	719	AAGACTATCATCATCTG--CTCCCTACAGGAGTGGAAACAGAGAGTGGCTGATTACATT	777
Qy	1195	-gccaaagatcatcttagaagaagaagaaaaagcatctgcagctgtgaaaag	1243
Db	778	AGCCACGGATCATTTAGCAAGAAGAACTAAGGCACTGTGAGCTTGGAACAG	827
RESULT	4		
BF686405			
LOCUS	781 bp	mRNA	22-DEC-2000
DEFINITION	602143786f1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304805 5',		
ACCESSION	BF686405		
VERSION	BF686405.1	GI:11971813	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		

DEFINITION	0202143786.F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304805 5'
ACCESSION	U02143786.F1
VERSION	BF086405
KEYWORDS	EST.
SOURCE	BF086405.1 GI:11971813
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapsfremail.nih.gov">cgapsfremail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CML170 row: 1 column: 22

High quality sequence stop: 731.

## FEATURES

Source

Location/Qualifiers

1..781  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4304805"  
/clone\_lib="NIH-MGC\_46"  
/tissue\_type="leiomyosarcoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pOH7; Site: 1; XhoI; Site-2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 235 a 168 c 204 g 174 t  
ORIGIN

Query Match 19.4%; Score 651.8; DB 11; Length 781;  
Best Local Similarity 99.4%; Pred. No. 1.3e-123;

Matches 675; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 385 ctgttgactcgggaattcgtgagccctcggagaaccggagccctctgccc 444  
D 5 cctgtgctgactcgggaattcgtgagccctcggagaaccggagccctctgccc 64  
QY 445 cccatcgaaccctcgaaccctcgaaccctcgaaccctcgaaccctcgaacc 504  
D 65 cctgtgctgactcgggaattcgtgagccctcggagaaccggagccctctgccc 124  
QY 505 gaatactcgaaccctcgaaccctcgaaccctcgaaccctcgaaccctcgaacc 564  
D 125 gaatactcgaaccctcgaaccctcgaaccctcgaaccctcgaaccctcgaacc 184  
QY 565 gaactcctggaatgaatgagagagagagagagagagagagagagagagag 624  
D 185 gacgtc-tgcataatgcatgagagagagagagagagagagagagagagagag 243  
QY 625 gctgtcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 684  
D 244 gctgtcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 303  
QY 685 cagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 744  
D 304 cagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 363  
QY 745 ctgtcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 804  
D 364 ctgtcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 423  
QY 805 tcaagaatgagatgagatgagatgagatgagatgagatgagatgagatgag 864  
D 424 tcaagaatgagatgagatgagatgagatgagatgagatgagatgagatgag 483  
QY 865 ctgagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 924  
D 484 ctgagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 543  
QY 925 tctgaatcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 984  
D 544 tctgaatcgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 603  
QY 985 agcttgagacataag-cacacatgagcagtgatcagagccatggagagatgag 1043  
D 604 agcttgagacataag-cacacatgagcagtgatcagagccatggagagatgag 663

QY 1044 tgaagaagaatgagcagca 1062  
D 664 tgaagaagaatgagcagca 682

## RESULT 5

BE882040

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE882040 1041 bp mRNA EST 20-OCT-2000  
601505326F2 NIH-MGC\_71 Homo sapiens cDNA clone IMAGE:3906774 5',  
mRNA sequence.  
BE882040.1 GI:10330816  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://img.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-f@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM9716 row: d column: 07  
High quality sequence stop: 586.

## FEATURES

Source

Location/Qualifiers

1..1041  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3906774"  
/clone\_lib="NIH-MGC\_71"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site-1: NotI; Site-2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb.  
BASE COUNT 385 a 258 c 223 g 175 t  
ORIGIN

Query Match 18.5%; Score 622.8; DB 11; Length 1041;  
Best Local Similarity 89.1%; Pred. No. 1.2e-117;

Matches 718; Conservative 0; Mismatches 82; Indels 6; Gaps 4;

QY 1923 ggaataaaagcggcaaaaagaagaagaagaagaagaagaagaagaagaaga 1982  
D 1 CGAATAAAAGCTCCCAAGAGCAAAATCCCAAGAGCGTTTCACAAACATTGAG 60  
QY 1983 gaactacaatgagagccctacaaatcaa-tgacacaaatgagatgagatgag 2041  
D 61 gaaactacaatgagagccctacaaatcaa-tgacacaaatgagatgagatgag 120  
QY 2042 atcttgaactcttataatgaagaagaagaagaagaagaagaagaagaaga 2101  
D 121 atcttgaactcttataatgaagaagaagaagaagaagaagaagaagaaga 180  
QY 2102 gtgagtgaagtgatgagatgagatgagatgagatgagatgagatgagatgag 2161  
D 181 gtgagtgaagtgatgagatgagatgagatgagatgagatgagatgagatgag 240  
QY 2162 aaccttgaagaatgagagaagaagaagaagaagaagaagaagaagaagaaga 2221  
D 241 aaccttgaagaatgagagaagaagaagaagaagaagaagaagaagaagaaga 300  
QY 2222 aatgttgaagaagctggctgaagaagaagaagaagaagaagaagaagaaga 2281





QY 3253 atgaagatctttaaatactatcaagtaacattatgatatgattatgatat 3312  
 |||  
 DB 116 AATGAAGATCTCTTTAAATAATCACTTAAACATTTAATATGATTATGATTATGAT 57  
 |||

QY 3313 tcatatgctacagagactgacataagaatcaataaatgattgttctactcg 3365  
 |||  
 DB 56 TCATATATGCTACAGAACTGACATTAAGATCAATTAATGATTGTTTACTCTG 4  
 |||

RESULT 8  
 A1806204/c 537 bp mRNA EST 19-DEC-1999  
 LOCUS w26c12.x1 Soares.NFL.T.GBC.S1 Homo sapiens cDNA clone  
 DEFINITION IMAGE:2356726 3', mRNA sequence.  
 ACCESSION A1806204  
 VERSION A1806204.1 GI:5392770  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 537)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert length: 729 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 444.  
 Location/Qualifiers  
 1..537  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2356726"  
 /clone\_1db="Soares.NFL.T.GBC.S1"  
 /lab\_host="DH10B"  
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site:1: Not 1; Site:2: Eco RI;  
 Equal amounts of plasmid DNA from three normalized  
 libraries (fetal lung MDHL19W, testis NHT, and B-cell  
 NCI-CGAP-GCB1) were mixed, and ss circles were made in  
 vitro. Following HAP purification, this DNA was used as  
 tracer in a subtractive hybridization reaction. The driver  
 was PCR-amplified cDNAs from pools of 5,000 clones made  
 from the same 3 libraries. The pools consisted of  
 1.M.A.G.E. clones 297480-302087, 682632-687239,  
 726408-728711, and 729096-731399. Subtraction by Bento  
 Soares and M. Fatima Bonaldo.  
 BASE COUNT 157 a 93 c 84 g 203 t  
 ORIGIN

Query Match 15.5%; Score 520.4; DB 10; Length 537;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-96;  
 Matches 532; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2832 caagagaatattgccaagcattacaagaataaccacataaacttcccttga 2891  
 |||  
 DB 537 CAAGAGAAATATTGCCAAGCATTACAGAAATTAACCCAT-ACATAATACCTTCCITGGAA 479  
 |||

QY 2892 aaacgcagctgctcagcctgtctcgggaagataatccatgtaattggaagaatgcatca 2951  
 |||  
 DB 478 AAACGCACTGTCTAGCTTACCTGTCTGGGAGACATATCCATGTAATTTGCAAAAATCATCA 419  
 |||

QY 2952 cgtcaataagaccagaattcaaggaacttacaattglaagagaagaacaaagcactgca 3011  
 |||  
 DB 418 CGTCAATATGACCCGAGAAITTCAGGAACCTTACATGTGTAGAGAAACAAACACTGCA 359  
 |||

QY 3012 aaagaagtgtcgcgaactcaataaataatgtaaaatcaltcgtgcaaaatggtcgaagcttg 3071  
 |||

DB 358 AAAGAACTGTGCGACATCATCAATAATATGATGAAATATCATCGAATAATGCGCAGGCTTG 299  
 |||

QY 3072 gggacaagatgagtgacacaaagccttagatttgccttgcctcaaaataagaaatltgt 3131  
 |||

DB 298 GGGAAACAATATATGATGTCGCAAAAGCTTAGATTGCTTCTGCTCAAAAATAAGAAATTTGT 239  
 |||

QY 3132 agtggctttcaaaataatccaagaagaacataacaaaagtggtaaatctat 3191  
 |||

DB 238 AGTGCTTTTCAAAAATATTCACAAAGAAACAAATACAAAAGCGGTACAAATTCCTAT 179  
 |||

QY 3192 cacatcccacatctgacattcaccattcagaatgctgttattagatgataagattacacat 3251  
 |||

DB 178 CACATTTCCCAATCTTACACTATTAGAAATGCTGTTTATTAGTATGAGCATTAAGCACTT 119  
 |||

QY 3252 gatgaagattcctttaaatactatcagttcaacattcaatgatatgattatgta 3311  
 |||

DB 118 GATTGAACATCTTTTAAATACATATCAGTTAAACATTATATGATTATGATTATGTA 59  
 |||

QY 3312 ttcattatgctacagaaactgacataagaatcaataaataatgatttctactcg 3365  
 |||

DB 58 TTCAATATGCTACAGAACTGACATTAAGATCAATTAATGATTGTTTACTCTG 5  
 |||

RESULT 9  
 A1922705/c 521 bp mRNA EST 01-SEP-1999  
 LOCUS w011f08.x1 NCI-CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2455047 3',  
 DEFINITION mRNA sequence.  
 ACCESSION A1922705  
 VERSION A1922705.1 GI:5658669  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 521)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Life Technologies catalog #: 11548-013  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdrr/image/image.html  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 410.  
 Location/Qualifiers  
 1..521  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2455047"  
 /clone\_1db="NCI-CGAP\_Pan1"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B"  
 /note="Organ: pancreas; Vector: PCMV-SPORT6; Site:1: SalI;  
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.72 kb. Life Technologies catalog #:  
 11548-013"

BASE COUNT 154 a 90 c 82 g 195 t  
 ORIGIN

Query Match 15.3%; Score 515.4; DB 10; Length 521;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-95;  
 Matches 516; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2849 agccttaagaataaacaccataacttcccttgcgaagaactgcaagtgtag 2908  
 |||

DB 521 AGCATTAACAAGAAATTAACCATCAATTAATACCTTCTTGCAGAAACGACAGTGTGCTAG 462  
 |||

QY 2909 cctgtctgggaagatataccatgttaattgagaataatgcatcactcaatataatgacccag 2968  
 |||||||  
 Db 461 CCTGTTTGGGGAAGATATCCATGTAATTGAGAAATATCATACGCTGATATGACCCAG 402  
 QY 2969 aattcaagaacttactacttgaagaagaacaaagacacgacgcaaaagaatgtgcgaact 3028  
 |||||||  
 Db 401 AATTCAGAGAACCTTACATTGTAAAGCAAAACAAACAAACCTGCAAAAGAGTGTCCGCACT 342  
 QY 3029 atcaaatatgtgtgaatcatctcgaatgtgccaatgtgccaatgtggaacaatgtatgtgc 3088  
 |||||||  
 Db 341 ATCAATATTAATGTGTAATTCATCTGCAATATGTGGCAGGCTTGGGGAACATGATGTGTC 282  
 QY 3089 acaaaagcttagatttgccttgcctcacaataaagaatcttgaatgtgttccaataa 3148  
 |||||||  
 Db 281 ACAAGGCTGTAATTTGCTGTCTGCAAAATAGGAATTTTGTAGTGTGTTTCAAAATA 222  
 QY 3149 attcaacaagaagaataatacaaaaagtggtgagaaattactatcacatttcccaacttg 3208  
 |||||||  
 Db 221 ATTCAACAAACAAACAAACAAACAAAGTGAGTAATTCATATGCAATTTCCCAATCTTG 162  
 QY 3209 actatcagaatgtcgttattatgtatgtgaggaatlaacacattgataagaattcttta 3268  
 |||||||  
 Db 161 ACTATTCGAATGCTGTTATTATTACTGATGAGGATAGCAGCTTGTATGAAATTTCTTTTA 102  
 QY 3269 aaataactatcagtaaacatttaataatataatgaatlaaagatcattcattactacagaa 3328  
 |||||||  
 Db 101 AATATCTATCATGATTAACATTTAATATGATTAATGATTAATGATTAATGATTAATGATTA 42  
 QY 3329 ctgacataagaactcaataaataatgattgttctactctg 3365  
 |||||||  
 Db 41 CTGACATTAAGATCAATTAATGATTTTACTCTG 5

RESULT 10  
 AA134958 552 bp mRNA EST 30-NOV-1996  
 LOCUS AA134958/c  
 DEFINITION z03104.s1 Striatagene colon (#937204) Homo sapiens cDNA clone  
 IMAGE:587743 3', mRNA sequence.  
 AA134958  
 AA134958.1 GI:1696041  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 552)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chippelli, B.,  
 Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,  
 M., Hultman, M., Kucada, T., Lacy, M., Le, N., Le, N., Mardis, E., Moore,  
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,  
 Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevas, E.,  
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)  
 97044478  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40M13 fwd, from Amersham  
 High quality sequence stop: 278.  
 Location/Qualifiers  
 1..552  
 /organism="Homo sapiens"  
 /db\_xref="GDB:4620045"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:587743"  
 /clone\_id="Striatagene colon (#937204)"

/tissue-type="tumor"  
 /cell\_line="T84 carcinoma cell line"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /note="Organ: colon. Vector: pBluescript SK-; Site: 1;  
 EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer:  
 01190 dt. T-84 colonic epithelial cell line. Average  
 insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor  
 sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'  
 CTCGACTTTTCTTTTCTTTT 3'

BASE COUNT 157 a 97 c 90 g 206 t 2 others  
 ORIGIN  
 Query Match 15.2%; Score 511.2; DB 10; Length 552;  
 Best Local Similarity 98.0%; Pred. No. 9.3e-95;  
 Matches 527; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 2828 aaacccagaagaatattgccaagcatttacaagaataaccacataaacttccctt 2887  
 |||||  
 Db 551 AAACCCAGGAAATATTGCCAGGCAAT-CCAGGATACCCCATCAATTAATCTTCTTT 493  
 QY 2888 gcaaaactgcaagtgctagccctgttctggaagaatatacctgtaattgaagaatgc 2947  
 |||||||  
 Db 492 GCAAAAGCTCAGTGTCTAGCCCTGTTCTGGGAAGATATCATGTAATTGCAAAATGC 433  
 QY 2948 atcacgcaatatgaccccaagaatcacaagaacttatacttgtaagaagaacaagcac 3007  
 |||||||  
 Db 432 ATCAGCTCAATATGACCCCAAGATTCAGGACTTTAATTGTAAGCAAAACAAACAC 373  
 QY 3008 tgcataaagaagtgctgcgactatcaaalaaatgtaagatcagaaatgtgcgaag 3067  
 |||||||  
 Db 372 TGCAAAAGAGGTGCGGACATATCAATTAATGCTGAAATCAATCTGCAATGTGGCCAG 313  
 QY 3068 ctgaggaacaatgatgatgacacaaagcttagatttgccttgcctcaaaataagaatt 3127  
 |||||||  
 Db 312 CTGGGGAACAAATGATGTGTCACAAAGGCTTGAATTTGCCCTTCTCAAAATAGGAATT 253  
 QY 3128 ttgtagtgatttcaaaaataatcacaagaagaacaatacaaaagtggtgagaaatc 3187  
 |||||||  
 Db 252 TTGTAGTGTTTCAAAATTAATTCACAAAGAACAAATGCAAAAGGTGGTGAATTAAC 193  
 QY 3188 ctatcacatttcccaactctgactatctcgaatgtcgttattatgtatgaatgac 3247  
 |||||||  
 Db 192 CTATGACATTTCCCAATCTGACATTCAGAAATGCTGTTAATGATGAGAGATTAGC 133  
 QY 3248 actgatgaagaattcttctttaaataactacatcagttaaacatttaataatgatttaa 3307  
 |||||||  
 Db 132 ACTTGATGGAAGATCTTTTAAATACTATCACTTAACATTTAATATGATTAATGA 73  
 QY 3308 tgtattcatatgtctacagaactgacataagaatcaataaataatgtatttactctg 3365  
 |||||||  
 Db 72 TGTATTCATTATGCTACAGAAAGTACATAGTAATCAATTAATGATGTTTACTCTG 15  
 RESULT 11  
 BE940626 546 bp mRNA EST 02-OCT-2000  
 LOCUS BE940626  
 DEFINITION RC3-UT0064-170800-021-e06 UT0064 Homo sapiens cDNA, mRNA sequence.  
 BE940626  
 BE940626.1 GI:10470130  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 546)  
 Nagai, M.A., Garcia, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, P.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.U.  
 Shotgun sequencing of the human transcriptome with ORF expressed



OY 3132 agtgggtttcaaaataaattcaacaagaagaacatacaaaaagtgggtaagaatactat 3191  
 |||||  
 DB 239 AGTGGTTCATAAAATATTACACCAAGCCCAATCCAAAGTGGGTAGAAATACCTAT 180  
 |||||  
 OY 3192 caactttcccaatttgcataattcaagaatgctgtttatttgaatgaagatagccct 3251  
 |||||  
 DB 179 CACATTTCACATCTTGCTAGTATTCAGAACTGCTGTTATTATTAAGTATAGCACTT 120  
 |||||  
 OY 3252 gattgaagaattcttcaaaatactacagttcaaacatttaataatgaatgaatga 3311  
 |||||  
 DB 119 AATTGAACATCTTTTAAATACATATCAGTTAAACATTATATGATTATGATTAACTGA 60  
 |||||  
 OY 3312 ttcatatgctacgaagactgacataagaatacaataaataatgatttactctg 3365  
 |||||  
 DB 59 TTCAATTATGCTACAGAACTGACATAGCAATCAATAAATGATTGTTACTCTG 6

RESULT 13  
 AM152541/c 484 bp mRNA EST 03-NOV-1999  
 DEFINITION x176b02.x1 NCI-CGAP Gas4 Homo sapiens cDNA clone IMAGE:2623947 3'

ACCESSION AM152541  
 VERSION AM152541.1 GI:6200441  
 KEYWORDS EST.

ORGANISM Homo sapiens  
 SOURCE human.

REFERENCE 1 (bases 1 to 484)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdrrp/image/image.html

Possible reversed clone: polyT not found  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 408.

FEATURES  
 source Location/Qualifiers

1..484  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:2623947"  
 /clone\_id="NCI-CGAP-Gas4"  
 /tissue\_type="poorly differentiated adenocarcinoma with  
 signet ring cell features"  
 /lab\_host="DH10B"  
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site: 1; Salt:  
 Site\_2: NotI; Cloned unidirectionally. Primer: 0190 dt.  
 Average insert size 1.69 Kb. Life Technologies catalog #:  
 11549-011"

BASE COUNT 145 a 88 c 74 g 177 t  
 ORIGIN

Query Match 14.3% Score 480.4; DB 10; Length 484;  
 Best Local Similarity 99.8% Pred. No. 1.9e-88;  
 Matches 481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2884 cttgcaaaactgagtgctagctgctgctggggaagatatacattgataatgagaaa 2943  
 |||||  
 DB 484 CTTGCAAAAACCTGACAGTGTAGCTGCTTCTGGGGAAGATATCCATGTAATTGAGAAA 425

OY 2944 atgcacacgcaalalagacccccagaatllcaagaacttatacttgaagagaacaaa 3003  
 |||||  
 DB 424 ATGCATCAGCTCATATATGACCCCAATTCAGAACTTACATTTAGACAAAAACAA 365  
 |||||  
 OY 3004 gcaactgcaaaagaagtgctgcagctatcaataaagtggaatctatcgaatgtgac 3063  
 |||||  
 DB 364 ACAGTCAAAAAGAAAGTGTGCGACTCAATTAATGAGTAAATCATCTGCAATATGTGC 305  
 |||||  
 OY 3064 cagcgttggggaacaatgagtgagcaagaagcttgaattgcttgccttcaaaaataag 3123  
 |||||  
 DB 304 CAGCGCTGGGGCAACATGATGATGCGCAAAAGCTTAGATTGCTGCTCAAAATAGG 245  
 |||||  
 OY 3124 aatttgtagtggtttcaaaaataaattcaacaagaacatacaaaaagtgggtaaga 3183  
 |||||  
 DB 244 AATTGTAGTGGTTCCTCAAAAATTAATTCACAAAGAACATTAACAAAAAGTGGTAGAA 185  
 |||||  
 OY 3184 ttaccatacaatttcccaatttgcataattcagaatgctgttatttgaatgagat 3243  
 |||||  
 DB 184 TTACCTATCAGATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTATGATGAGAGAT 125  
 |||||  
 OY 3244 taagacttgatgaagattcttcaaaatactatacaagtaaacatttaataatataga 3303  
 |||||  
 DB 124 TAGCAGCTTGATTCAGAGTTCTTTAAATATCTATCACTTAAACATTTAATATGATATGA 65  
 |||||  
 OY 3304 ttaatgattcattatgctacagaactgacataagaatacaataaataatgatttact 3363  
 |||||  
 DB 64 TTAATGATTATCATATTATCTACAGACTGACATAGCAATCAATAAATGATTGTTACTC 5  
 |||||  
 OY 3364 tg 3365  
 |||||  
 DB 4 TG 3

RESULT 14  
 AM589567/c 475 bp mRNA EST 22-MAR-2000  
 LOCUS x012b12.x1 NCI-CGAP\_Ut3 Homo sapiens cDNA clone IMAGE:2703743 3'

DEFINITION similar to TR:095786 095786 RNA HELICASE. ; mRNA sequence.  
 ACCESSION AM589567  
 VERSION AM589567.1 GI:7276681  
 KEYWORDS EST.

ORGANISM Homo sapiens  
 SOURCE human.

REFERENCE 1 (bases 1 to 475)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 Image.llnl.gov/image/html/iresources.shtml

Possible reversed clone: polyT not found  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 407.

FEATURES  
 source Location/Qualifiers

1..475  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:2703743"  
 /clone\_id="NCI-CGAP-Ut3"  
 /tissue\_type="poorly differentiated endometrial  
 adenocarcinoma, 2 pooled tumors"  
 /lab\_host="DH10B"





.....

---